

SEQUENCE LISTING

<110> BENOIT, Patrick
SCHWARTZ, Bertrand
BRANELLEC, Didier
CHIEN, Kenneth R.

<120> SEQUENCES UPSTREAM OF THE CARP GENE, VECTORS CONTAINING
THEM AND USES THEREOF

<130> 03806.0530-00000

<140>

<141> 2001-12-07

<160> 5

<170> PatentIn Ver. 2.1

<210> 1

<211> 2358

<212> DNA

<213> Mus musculus

<400> 1

```

ggatcctttc atgtttaaca atatcaacc taaccaagg ggaacagcct gcctgacagt 60
ggctttgccca cccatgaata cttcctagtc tagtccgttt gtgaaactca gcccatccca 120
acacttctgc aagccccatc ctctacaagg tgctcattgg gaatttcctg gagcttctct 180
ttcaggatca gcctgattct agggcagcag ttctcaacct gggggcctcg accccttttg 240
gggaatcaaa cgacccttta caggggtcac atatcatcta tcctatatgt caggtattta 300
cattacgatt cgtaacagta gcaaaattac aggtatgaaa tagcaatgaa ataattttat 360
gattgaaggt caccacaaca tgaggccgcc acactgttct agagaaaaat cacctgggtg 420
gggaaagggt tgggaaagcc tttctgtcca ttcttcattc ttcaaagtga tgtgttcaca 480
gaaagccttt cagctgttct gctggggctc ttagtaagtc tgagtaggaa ctgtatgtac 540
caggtctgct tcttatgggt ggagccaaga cgcacgtggt gtggagcgaa gacgcaacct 600
caccttctag ctctgcatcc atagcaagta gcctaattgt tctgtgtcta ggtgtcatct 660
ctgtgaatcg agatccttgg ccttgcttga attagggagg cacaaaatac tcagagattc 720

```

aagactgctc agcagcccag agtccttcct caaaggaaag gtctcaactc tcagcccccc 780
ttagctctga gtcaggcctg gaacaaacgg ccacaggaat gagaaaagct gccatagctg 840
cttgtcactt caagagggtca aagaaaatag tgттаaccat gaaaacgaga agaccaacag 900
ttatccattg atagcgtctc aggacagata ggacagagag aacactagga gaggggaacc 960
cacgaaggac aaggtattag tgtgttggtt ttcagggcaa tgtcttgtag tgaagattct 1020
agaaacacaa tttgctgggt gaacagctga agtgggggtg gggttcttac cccatgttca 1080
tggaaggggtg agtgaggaga gacagatata tgatggccag cataacaaac atacacaaca 1140
ccctaattaa cacttcctc ttctactgac accccttca ctctcctctt tcataaaaaa 1200
taaaaaaagt attttatgtg gctcttacga tagaatcttt cctcgaacta taaaagatc 1260
taaataattta tatttttcac attttaatat cttagcgatg acaagccaga aacaagtatt 1320
ttttgcctct ctcaacagca aagcttgggg cttttttgtt tccgtgtag gaatagaaca 1380
cgagagcccc gtgtatctag gcagatgctc tatcattagc ccatgagtct ccagcctcag 1440
acgcacattt ttctcgggct ctcttaagct tttccacag cattgggaaa ctttactgac 1500
agcatccaag ttgtgcttct gctaagaact ggactcacat ctctctgtgc atcacttcgg 1560
cccgtttttg ggtagatcct ctgattagcc ttcagattta gaacacggtg agcctgtggt 1620
gcactaatta tggccagtga caccatagag tcaaagtgc ttactgaatg ctttcaattt 1680
ctcctaatgc tggtacgatg gcatgtcaca gggccatttt agctgcagac atcactccag 1740
agaattccaa acagatagag acaagtggca cccagacca tctccttccc ctcgggctga 1800
ttatccccag aaataggatg tcccaaagca acacttcca gccaactgga gtgctgataa 1860
gtccagttat cagaaagata tggctgtaag tgtgatgcac agtgcttgca ttttcttgat 1920
acgtagtca tatgagagct gacaaagaag gaaaaagagc agcgatgtgg tgcaatatta 1980
acaggcagct gtcccctggc ttcccgatac gtgggatgac tcgcattgct gagcgggtgtg 2040
gtcactgcc aaggaatgac cctctcacat ttcttcctga ttcgcatacg ccgcggccag 2100
cttgtcatct ccctcttggg cttcccagac actaagtctg gaatgaaaat tcacctgcct 2160
ctgaattggc cactgggtggg ggcaggggtg tgacttggct tcccaggctg gaagattatc 2220
tcaccagcc ctagctatat aacgggctgg tgtggagggg ctccacaggg ccagttccag 2280
gggttcatcc acaagagaga aaaacataga ctcgaggtct agggagcttg catgcctgca 2340

ggtcggaggc caccatgg

2358

<210> 2

<211> 2074

<212> DNA

<213> Homo sapiens

<400> 2

ctgcagcaag ttacttaatg ttttttgcct cagcatcctc tctgtaaaat gagagcatta 60
gtcttgctcc aacttcgagg gcatggacag ctctgggatt tcatatccaa gacccttaaa 120
catcccacag tccttcccc aaacacttct cctcctaata cctccctcag tttgggtcag 180
gcctggaaca aaaaggcata cgaaatggta gaaaaagtgt ccatgactac ttctgactta 240
gatgaagaga ccaatgaaaa tagtaatgac tctgtttgct tcagcaggac atatactaaa 300
ataggagcta tacaagaag attagcatgg actctgtgca agaatgacac acaaatttgt 360
gaaacattcc atatattaaa aataaataaa taataaagag aaaaggaaaa aattaaaaag 420
aaaatagtga tagctgtgtc catctcaaag aaaagcccag gagatttcct ttattttacc 480
cctttaagat agaatattag gagaccggaa catatgatac aggaggtact gggaggggtcc 540
ctctttgtca atgttttgtc ttgggggtggg gagtcgatgt cttctcaaag tttcagaaac 600
accatccact gactgagcat tcaaggggca agaggagaat ggcagccaca tttgttgatt 660
gggtgagttt ggggagaaat agacacacaa aggtcaaaca taacttccta attaacactt 720
ccctccattc acaattccct tctccattc ttctctctg tcttttacts akaraaacc 780
agtttttcct gaaactataa aaataccccc agtatgttta cataatttac acctcaaaga 840
ttagaaacca gaaatagaga ccttttcaac ctttccggaa gcaaagtgca ttatccctcc 900
agccacgtgt ctcaaattct gatgcatcag aatcatctgg gtgctttkaa attcaagatg 960
attcctacga gttaccataa atcaactcag aattccctgg agtggggcca gggatctgta 1020
tttctgacaa gctcccacag gtgattcctt tccccacagc atttgagaac ttcagctcaa 1080
tgacctaatc agagtcctgc cattgctaata atctgggtctc atttttbtca tatatatata 1140
tagtatttgt ggtagagatg ggattttgcc atgttgccca ggctagtatt gaactcctaa 1200
gctaagcaat cttcctgtct ctgcctccca aaatgttggg attacaggtg taagccactg 1260



cacccggctg atagctggtt tcatttactc tatttcttga ccactctgat ccattttgaa 1320
gtaaaaatgc tccaattatt atgctgtttt agaacacggt aagcatgtca tgtgctaata 1380
gccagtgaca tcataaaaga aaagtgcatt actgaatgct ttcaatgtct tataatgatg 1440
gtaaggtggc atgtcatggg gcctatttag ccagacatc actccaaaga attcceaaca 1500
gatatagaca agtgccttta gggcccagat cccttcccct caggctgttt acccagggaa 1560
taggatgtcc tgggacaagt ttcccctaag tgaagtgttg ataagtctgc ttatcagaaa 1620
gatattactg ggggtgtgat atgtagggca tctacatttt cttgatagggt agtcatatga 1680
aagctgacaa agaaaaaaag ggcagtgatg tggtgcaatg tcaacagaca gctgtcccct 1740
gactcttgac aaataggatg acttgcattg ctgagcgatg tgatcaccac caaaggaatg 1800
gccctctcac atttcttcct gattcacata ttcagcaggg ttagcttgtc ctcccctccc 1860
tcttcagctt ccagacact gagtctggaa tgaaaattca cctgcctctg agttggctcc 1920
taatgggggc gggagtgtta cttcggttcc caggttggaa gattatctca cccggcccca 1980
gctatataag ctgaccggtg tggagggggcc cagcagggcc aactccaggg attccttcca 2040
cgacagaaaa acatacaaga ctccttcagc caac 2074

<210> 3
<211> 750
<212> PRT
<213> Homo sapiens

<400> 3
Met Gly Glu Thr Leu Gly Asp Ser Pro Ile Asp Pro Glu Ser Asp Ser
1 5 10 15
Phe Thr Asp Thr Leu Ser Ala Asn Ile Ser Gln Glu Met Thr Met Val
20 25 30
Asp Thr Glu Met Pro Phe Trp Pro Thr Asn Phe Gly Ile Ser Ser Val
35 40 45
Asp Leu Ser Val Met Glu Asp His Ser His Ser Phe Asp Ile Lys Pro
50 55 60
Phe Thr Thr Val Asp Phe Ser Ser Ile Ser Thr Pro His Tyr Glu Asp
65 70 75 80
Ile Pro Phe Thr Arg Thr Asp Pro Val Val Ala Asp Tyr Lys Tyr Asp
85 90 95

Leu	Lys	Leu	Gln 100	Glu	Tyr	Gln	Ser	Ala 105	Ile	Lys	Val	Glu	Pro 110	Ala	Ser		
Pro	Pro	Tyr 115	Tyr	Ser	Glu	Lys	Thr 120	Gln	Leu	Tyr	Asn	Lys 125	Pro	His	Glu		
Glu	Pro 130	Ser	Asn	Ser	Leu	Met 135	Ala	Ile	Glu	Cys	Arg 140	Val	Cys	Gly	Asp		
Lys 145	Ala	Ser	Gly	Phe	His 150	Tyr	Gly	Val	His	Ala 155	Cys	Glu	Gly	Cys	Lys 160		
Gly	Phe	Phe	Arg	Arg 165	Thr	Ile	Arg	Leu	Lys 170	Leu	Ile	Tyr	Asp	Arg 175	Cys		
Asp	Leu	Asn	Cys 180	Arg	Ile	His	Lys	Lys 185	Ser	Arg	Asn	Lys	Cys 190	Gln	Tyr		
Cys	Arg	Phe 195	Gln	Lys	Cys	Leu	Ala 200	Val	Gly	Met	Ser	His 205	Asn	Ala	Ile		
Arg	Phe 210	Gly	Arg	Met	Pro	Gln 215	Ala	Glu	Lys	Glu	Lys 220	Leu	Leu	Ala	Glu		
Ile 225	Ser	Ser	Asp	Ile	Asp 230	Gln	Leu	Asn	Pro	Glu 235	Ser	Ala	Asp	Leu	Arg 240		
Ala	Leu	Ala	Lys	His 245	Leu	Tyr	Asp	Ser	Tyr 250	Ile	Lys	Ser	Phe	Pro 255	Leu		
Thr	Lys	Ala	Lys 260	Ala	Arg	Ala	Ile	Leu 265	Thr	Gly	Lys	Thr	Thr 270	Asp	Lys		
Ser	Pro	Phe 275	Val	Ile	Tyr	Asp	Met 280	Asn	Ser	Leu	Met	Met 285	Gly	Glu	Asp		
Lys	Ile 290	Lys	Phe	Lys	His	Ile 295	Thr	Pro	Leu	Gln	Glu 300	Gln	Ser	Lys	Glu		
Val 305	Ala	Ile	Arg	Ile	Phe 310	Gln	Gly	Cys	Gln	Phe 315	Arg	Ser	Val	Glu	Ala 320		
Val	Gln	Glu	Ile	Thr 325	Glu	Tyr	Ala	Lys	Ser 330	Ile	Pro	Gly	Phe	Val 335	Asn		
Leu	Asp	Leu	Asn 340	Asp	Gln	Val	Thr	Leu 345	Leu	Lys	Tyr	Gly	Val 350	His	Glu		
Ile	Ile	Tyr 355	Thr	Met	Leu	Ala	Ser 360	Leu	Met	Asn	Lys	Asp 365	Gly	Val	Leu		
Ile	Ser	Glu	Gly	Gln	Gly	Phe	Met	Thr	Arg	Glu	Phe	Leu	Lys	Ser	Leu		

370					375					380					
Arg 385	Lys	Pro	Phe	Gly	Asp 390	Phe	Met	Glu	Pro	Lys 395	Phe	Glu	Phe	Ala	Val 400
Lys	Phe	Asn	Ala	Leu 405	Glu	Leu	Asp	Asp	Ser 410	Asp	Leu	Ala	Ile	Phe 415	Ile
Ala	Val	Ile	Ile 420	Leu	Ser	Gly	Asp	Arg 425	Pro	Gly	Leu	Leu	Asn 430	Val	Lys
Pro	Ile	Glu 435	Asp	Ile	Gln	Asp	Asn 440	Leu	Leu	Gln	Ala	Leu 445	Glu	Leu	Gln
Leu	Lys 450	Leu	Asn	His	Pro	Glu 455	Ser	Ser	Gln	Leu	Phe 460	Ala	Lys	Leu	Leu
Gln 465	Lys	Met	Thr	Asp	Leu 470	Arg	Gln	Ile	Val	Thr 475	Glu	His	Val	Gln	Leu 480
Leu	Gln	Val	Ile	Lys 485	Lys	Thr	Glu	Thr	Asp 490	Met	Ser	Leu	His	Pro	Leu
Leu	Gln	Glu	Ile 500	Tyr	Lys	Asp	Leu	Tyr 505	Ala	Trp	Ala	Ile	Leu 510	Thr	Gly
Lys	Thr	Thr 515	Asp	Lys	Ser	Pro	Phe 520	Val	Ile	Tyr	Asp	Met 525	Asn	Ser	Leu
Met 530	Met	Gly	Glu	Asp	Lys	Ile 535	Lys	Phe	Lys	His	Ile 540	Thr	Pro	Leu	Gln
Glu 545	Gln	Ser	Lys	Glu	Val 550	Ala	Ile	Arg	Ile	Phe 555	Gln	Gly	Cys	Gln	Phe 560
Arg	Ser	Val	Glu	Ala 565	Val	Gln	Glu	Ile	Thr 570	Glu	Tyr	Ala	Lys	Ser 575	Ile
Pro	Gly	Phe	Val 580	Asn	Leu	Asp	Leu	Asn 585	Asp	Gln	Val	Thr	Leu 590	Leu	Lys
Tyr	Gly	Val 595	His	Glu	Ile	Ile	Tyr 600	Thr	Met	Leu	Ala	Ser 605	Leu	Met	Asn
Lys	Asp 610	Gly	Val	Leu	Ile	Ser 615	Glu	Gly	Gln	Gly	Phe 620	Met	Thr	Arg	Glu
Phe 625	Leu	Lys	Ser	Leu	Arg 630	Lys	Pro	Phe	Gly	Asp 635	Phe	Met	Glu	Pro	Lys 640
Phe	Glu	Phe	Ala	Val 645	Lys	Phe	Asn	Ala	Leu 650	Glu	Leu	Asp	Asp	Ser 655	Asp

Leu Ala Ile Phe Ile Ala Val Ile Ile Leu Ser Gly Asp Arg Pro Gly
 660 665 670
 Leu Leu Asn Val Lys Pro Ile Glu Asp Ile Gln Asp Asn Leu Leu Gln
 675 680 685
 Ala Leu Glu Leu Gln Leu Lys Leu Asn His Pro Glu Ser Ser Gln Leu
 690 695 700
 Phe Ala Lys Leu Leu Gln Lys Met Thr Asp Leu Arg Gln Ile Val Thr
 705 710 715 720
 Glu His Val Gln Leu Leu Gln Val Ile Lys Lys Thr Glu Thr Asp Met
 725 730 735
 Ser Leu His Pro Leu Leu Gln Glu Ile Tyr Lys Asp Leu Tyr
 740 745 750

<210> 4
 <211> 30
 <212> DNA
 <213> Artificial sequence
 <400> 4
 GGCGATTAA ATAATGTAGT CTTATGCAAT

30

<210> 5
 <211> 31
 <212> DNA
 <213> Artificial sequence
 <400> 5
 GGGGTCTAGA AGGTGCACAC CAATGTGGTG A

31